"STATISTICS" IN A MATHEMATICAL ENCYCLOPEDIC DICTIONARY.

By H. L. RIETZ, State University of Iowa.

In commenting on Professor G. A. Miller's article in this MONTHLY (*1918*, 383) giving meanings of *group* and *group theory*, Professor E. R. Hedrick, Chairman of the Association's Dictionary Committee, suggested (*1918*, 428) that further sample definitions be submitted. As it is not unlikely that there will be considerable difficulty in determining the extent to which terms from applied mathematics should be included in the proposed dictionary, the writer has been considering the question with special reference to the mathematics of statistics, probability, and insurance. As a result of such consideration, the following brief paper is submitted to explain the meanings of *statistics* and of the associated expressions,[1] *statistical data, statistical methods, theory of statistics, mathematical statistics, statistical probability,* and to suggest a list of terms and expressions from statistics, probability, and insurance that should probably be included in the dictionary. In the selection of the list of terms and expressions given below, the writer has been guided by his experience with seniors and first-year graduate students taking courses in statistics and actuarial theory, and has included only terms and expressions within the range of reading of such students.

Statistics (stă-tĭs'tiks), F. *statistique*, G. *Statistik*, I. *Statistica*, Sp. *estadistica*. The word statistics seems to be derived from the Latin *status*, used in the sense of a political state. Statistics is a comparatively new word. Its first occurrence in English thus far noted seems to be in J. F. von Bielfeld, *The Elements of Universal Erudition*, translated by W. Hooper, London, 1770. One of the chapters of this book is called *Statistics*, and the subject is defined as "The science that teaches us what is the political arrangement of all the modern states of the known world." The word occurs in 1787 with a somewhat changed meaning in the preface to E. A. W. Zimmerman, *A Political Survey of the Present State of Europe*. In this work it is stated that about forty years before the branch known as statistics was formed into a separate science in Germany. The German word *statistik* was used by Professor Achenwall of Göttingen in 1749, and the Latin, *statisticus*, was used at a somewhat earlier date. In *Meyers Konversationslexikon*, 6 ed., volume 18, under "Statistik", Schäzer (1735–1809), a pupil of Achenwall, defined: "Statistik ist stillstehende Geschichte; Geschichte ist fortlaufende Statistik."[2] In 1790 Sir John Sinclair stated in a letter to the Clergy of the Church of Scotland that "statistical inquiries" have been carried to a great extent in Germany, and adds that the expression "statistical inquiries" means "inquiries respecting the population, the political circumstances, the productions of a country, and other matters of state." *Statistics*, as thus used by German writers and by others in the eighteenth century, meant an exposition of the character of the state, and such expositions were usually verbal rather than numerical. With the growth of official numerical data, it was natural that numerical statements should begin to replace the verbal statements. Statistics thus gradually came to mean an exposition of the attributes of the state by numerical methods. Following this usage, the word next came to denote the figures used in such descriptions. Thus, the collections of numerical data were called statistics. This use of the word prevails at the present time but the data may refer to the state or to any other subject.

There is, however, an element in the meaning of the word statistics as used at the present time in the theory of statistics that is not necessarily involved in a collection of figures. Thus, a collection of 1000 numbers consisting

[1] These expressions should be listed alphabetically and references should be given to see their meanings under *statistics*.

[2] This citation is contributed by Professor A. J. Kempner.

of the number 5 written 1000 times would not constitute statistics. Numerical data known as statistics or *statistical data* have a certain element of variability. For example, statistics on the statures of men are variable from man to man. Statistics of social interest show great variability from individual to individual and from community to community. Statistics of meteorology show great variability from time to time and from place to place. The term statistics as used at present in the theory of statistics means numerical data that exhibit variability in individual items, where such variability is ascribed to a multiplicity of causes. See G. U. Yule, *Introduction to the Theory of Statistics*, London, 1922, pp. 1–5; H. Bruns, *Wahrscheinlichkeitsrechnung und Kollektivmasslehre*, Leipzig and Berlin, 1906, pp. 1–17; P. M. H. Laurent, *Statistique Mathématique*, Paris, 1908, pp. 1–16; and E. Blaschke, *Vorlesungen über mathematische Statistik*, Leipzig and Berlin, 1906, pp. 1–8.

The expression *statistical methods* means methods which are suitable for the description and characterization of statistical data. In the development of the meaning of statistical methods there have been influences in operation from three sources—the calculus of probability, the preparation of life and monetary tables under the name "political arithmetic," and the collection of data to be used in the machinery of government and business.

An exposition of the principles on which statistical methods are based is called the *theory of statistics*. A set of mathematical propositions that relate to statistical methods is often called *mathematical statistics*.

The relation of mathematical statistics to the theory of probability may be indicated by saying that the general problem of mathematical statistics in its ideal form is to determine a system of drawings to be carried out with urns of fixed composition, in such a way that the results of the set of drawings lead, with a high degree of probability, to a table of values identical with the statistical data (cf. E. Borel, *Éléments de la Théorie des Probabilités*, Paris, 1910, p. 167). Mathematical statistics is thus one branch of the theory of probability (cf. articles on *statistics* in German and French encyclopedias of mathematics). In fact, *a posteriori* or inductive probability is sometimes called *statistical probability*. (See E. Czuber, *Wahrscheinlichkeitsrechnung*, volume 2, Leipzig and Berlin, 1921, p. 6.) The concept of statistical probability is involved whenever the properties of an aggregate are predicted or inferred by observation of a sample taken from the aggregate. Many such inferences are drawn by persons unfamiliar with mathematical statistics, and there is practically no doubt that many conclusions thus obtained are invalid.

Mathematical statistics aims to establish criteria that give numerical values to the degrees of confidence to be placed in such inferences. In the development of these criteria statistical probability means the limiting value, as s becomes infinite, of m/s, where m is the frequency of happening of the event in s trials. The existence of the limit is assumed. The applications of mathematical statistics cover a wide range of scientific and social interests. These applications include the whole theory of insurance, and have an important place in biology, anthropology, psychology, economics, and even in the more exact sciences of chemistry and physics. To give a notion of the variety of applications, we may cite the following: *Biometrika*, "a journal for the statistical study of biological problems"; E. L. Thorndike, *Educational Psychology*, New York, 1913–1914; H. L. Rietz and H. H. Mitchell, "On the metabolism experiment as a statistical problem" (*Journal of Biological Chemistry*, volume 8, 1910, pp. 297–326); E. Rutherford and H. Geiger, "The probability variations in the distribution of α particles" (*Philosophical Magazine*, series 6, volume 26, 1910, pp. 698–707); J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, New York, 1902; F. Y. Edgeworth, "On the application of probabilities to the movement of gas-molecules" (*Philosophical Magazine*, series 6, volume 40, 1920, pp. 249–272). Lists of references on the methods and theory of statistics are given in G. U. Yule, *Introduction to the Theory of Statistics*, London, 1922, and E. Blaschke, *Vorlesungen über mathematische Statistik*, Leipzig and Berlin, 1906. A bibliography of applications prior to 1904 is given in C. B. Davenport, *Statistical Methods*, New York, 1904.

The following terms and expressions drawn from the theories of probability, statistics, and insurance are tentatively suggested for inclusion in the proposed dictionary. It seems that any term or expression that occurs infrequently, and that would require an appreciable amount of space to give its meaning in the dictionary, may well be treated by giving one or more references to its definitions

and use in the literature. In accord with this view, the mark "ref." is placed after each term that may, in the judgment of the writer, be treated in this way. Doubtless a considerable number of other terms may be treated by giving a very brief statement and references. In the preparation of the following list, the writer is indebted to Professor E. L. Dodd for valuable suggestions and additions.

Accumulation of discount, ref.
Actuarial theory
Adjustment of data, ref.
Advowson, ref.
Aggregate mortality table
Allocurtic, ref.
Amortization
Amortization of premium, ref.
Annual rent
Annuity,
  apportionable
  certain
  complete
  continuous
  deferred
  due
  forborne
  immediate
  intercepted
  joint life
  last survivorship
  life
  perpetual
  reversionary
  survivorship
  temporary
A posteriori probability
A priori probability
Array
Arithmetical mean
Association, theory of, ref.
Assurance
Automatic policy loan
Average
Average deviation
Bayes rule, ref.
Benefit of insurance, ref.
Bernoulli series, ref.
Bernoulli theorem, ref.
Bias in sampling
Biometry
Capitalized cost
Charlier Coefficient of Disturbancy
Claims, death
Class
Class frequency
Class interval
Class mark
Coefficient of association
Coefficient of contingency, ref.
Coefficient of variability
Commutation columns
Commutation symbols
Complement of life, ref.
Complete expectation of life
Compound reversionary addition, ref.
Contingency coefficient, ref.
Continuous instalment
Convertible term
Copyhold, ref.
Correlation,
  multiple, ref.
  normal
  partial
  rank
  spurious
Correlation coefficient
Correlation ratio
Cost of insurance
Cumulative graphs
Curtate expectation of life
Death rate, central
Death strain
Decile
Dependent events
Differences,
  central
  finite
  order of
Discount
Dispersion
Dividend (in life insurance),
  annual
  deferred
Endowment, pure
Endowment insurance
Equally likely
Equated time
Equation of life,
  of payments
  of value
Error
Errors, theory of
Expectation of life,
  complete
  curtate
Expiry
Extended insurance

Extrapolation
Figurate numbers
Fluctuations in sampling
Force of discount
Force of interest
Force of mortality
Frequency curve
Frequency distribution,
    binomial
    Gaussian
    normal
    skew
Frequency polygon
Frequency surface
Gain and loss exhibit
Gauss curve of error
Geometrical mean
Geometrical probability
Gompertz's law, ref.
Goodness of fit
Graduation of data
Group insurance
Hardy's formula, ref.
Harmonical mean
Heteroclitic, ref.
Heterograde series, ref.
Heteroscedastic, ref.
Heteroscedasticity, ref.
Histogram, ref.
Historigram, ref.
Homoclitic, ref.
Homogeneity in statistics
Homograde series, ref.
Homoscedastic, ref.
Homoscedasticity, ref.
Homotyposis, ref.
Incontestible
Independent events
Index of abmodality
Index number
Industrial insurance
Initial expense
Insurable interest
Insurance,
    assessment
    capital redemption
    casualty
    fraternal
    life
    property
Interest,
    accumulative rate of
    compound
    continuously convertible
    effective rate of
    frequency of conversion of
    instantaneous rate of
    nominal rate of
    rate of
    remunerative rate of
    simple
International actuarial notation, ref.
Isocurtic, ref.
Isotropic, ref.
Joint life annuity
Joint life insurance
Joint life probability
Lag
Lapse
Last survivorship annuity, ref.
Last survivorship insurance, ref.
Least squares, ref.
Legal reserve insurance
Level premiums
Lexis ratio, ref.
Lexis scheme, ref.
Life interest
Lifetime, most probable
Limited payment policy
Loading
Logarithmic paper, ref.
Lorenz graph
Lubbock's formula, ref.
Makeham's laws
Mathematical risk, ref.
Mathematics of statistics
Maturity
Mean contingency, ref.
Mean deviation
Mean error
Mean square contingency, ref.
Mean value
Median
Method of moments, ref.
Mode
Modulus as a measure of dispersion
Moments
Mortality rate,
    instantaneous
Mortality table,
    aggregate
    select
    ultimate
Moving average
Mutuality
Natural premium
Net premium
Net valuation
Nominal interest
Non-forfeiture
Normal frequency curve
Normal law of error
Normal probability curve
Ogive
Old line insurance
Ordinary life policy

Paid up insurance
Panmixia
Participating policies
Percentile
Perpetuity
Poisson scheme, ref.
Poisson series, ref.
Poisson theorem, ref.
Policy,
  continuous instalment
  endowment
  limited payment
  ordinary life
  pure endowment
  term
Polychroic functions, ref.
Precision
Premium,
  annual
  gross or office
  natural
  net
  single
Probability,
  a posteriori
  a priori
  deductive
  empirical
  inductive
  statistical
  theory of
Probable error
Probable lifetime
Prospective method
Pure endowment
Quartile
Quartile deviation
Radix of a table
Random sampling
Range
Rank
Rating up
Regression, linearity of
Regression coefficient
Regression curve
Reinsurance
Renewable term insurance
Reserve,
  initial
  mean
  modified preliminary term
  net level premium
  preliminary term
  select and ultimate
  terminal
Retrospective method
Reversion
Reversionary annuity
Reversionary expectations of life, ref.
Rests
Root-mean-square-deviation
Sampling
Schedules
Sheppard's corrections
Simple and compound survivorship, ref.
Simple sampling
Sinking fund
Skew frequency curve
Skewness
Smoothing
Spearman's coefficient, ref.
Square root of mean square
Standard deviation
Statistical mechanics
Statistical methods
Statistical probability
Statistics,
  theory of
  vital
Stirling's theorem, ref.
Surplus of an insurance company
Surrender charge
Surrender value
Tchebycheff's theorem
Term policy
Term premium
Tetrachroic functions, ref.
Tontine
Total and permanent disability
Trend
Uniform seniority
Valuation of policies, see reserve
Valuation date of issue
Variant
Variate,
  graduated
  integral
Vie probable
Weighted arithmetical mean
Woolhouse's formula, ref.
Yule coefficient, ref.